Canada. Statistics bureau
Analytical and technical memorandum
No. 3. The geographic basis of the DBS geocoding system for urban areas: an overview. 1969.

CANADA
DOMINION BUREAU OF STATISTICS
CENSUS DIVISION

ANALYTICAL AND TECHNICAL MEMORANDUM

No. 3

THE GEOGRAPHIC BASIS OF THE DBS GEOCODING SYSTEM FOR URBAN AREAS: AN OVERVIEW

by

Robert J. Ion
Geography Section

Ottawa, December 1969

TABLE OF CONTENTS

LIST OF FIGURES

THE GEOGRAPHIC BASIS OF THE DBS GEOCODING SYSTEM FOR URBAN AREAS: AN OVERVIEW*

Résumé

L'auteur examine la base géographique du système de géocodage urbain du B.F.S. en vue de souligner l'importance du cadre spatial choisi pour la réalisation de l'objectif du système qui est de permettre la tabulation des données statistiques pour les régions délimitées 'ad hoc' par l'utilisateur. L'étude montre que trois éléments principaux du cadre spatial: le côté d'îlot, sa série d'adresses et ses coordonnées géographiques, exercent une forte influence sur chacune des phases suivantes du système: le géocodage, le stockage des données et leur extraction. L'auteur en conclut que le développement de n'importe quel système d'information spatiale nécessite une recherche approfondie des éléments géographiques qui en fin de compte en contrôlent l'usage.

Abstract

The geographic basis of the DBS urban geocoding system is examined with a view to underlining the importance of the system's spatial framework in the attainment of its objective; namely, special tabulation of data by 'ad hoc' user-specified areas. The study indicates that three main elements of the spatial framework — the block-face, its address range and its geographic co-ordinates — exert a strong influence on the system's geocoding, data storage and data retrieval stages. The author concludes that the development of any spatial information system necessitates research into the geographic elements that ultimately control system utility.

---

* This is a revised version of a paper drafted in November 1968.
The author is grateful to his colleagues on the geocoding system team and elsewhere in DBS for their critical reading of the final draft and for their helpful comments. The author is, however, solely responsible for any errors or deficiencies in this paper.

1. Introduction

1.1 The General Context

Over the past two years the Dominion Bureau of Statistics (DBS) has conducted research into defining the small-area information problem and into developing a computerized system for its solution and for general utility. The outcome of this work is the Geographically Referenced Data Storage and Retrieval System, referred to throughout this paper as the DBS geocoding system. Two variants of the system are being developed, one for larger urban areas and the other one, a compatible approach, for the remaining area of Canada. In this paper only the system for urban areas is discussed, and the term 'urban' in this context connotes an area suited to the existing street address conversion mode of geocoding.

The small area information problem may be defined briefly, and in a geographical context, as a burgeoning demand for data pertaining to micro-areas; areas of sub-municipality size, bearing little or no relationship to existing statistical area units, but nonetheless significant in the decision-making processes of management and planning, and in general conducive to applied research in the social and economic fields.(1) Increasingly, the spatial attribute of data observations has been imposing a constraint upon the general usefulness of data. Haggett (1965) enunciated this constraint on data utility and referred to the problem of "the yoking of locational observations to the characteristics of the collecting area". Hagerstrand in Sweden sought as early as 1955 "to investigate if a possibility exists to give to the space-aspect of statistical material the same neutrally objective status as the time-aspect always had" (Hagerstrand, 1955). Another problem with small-area information arises from the fact that much information is coded to areas such as census tracts often in a non-geographic or quasi-geographic format, thus making it impossible to

(1) Further insight may be obtained from the series Census Tract Conference Papers, Bureau of the Census, Washington, D.C. See also Simmons, 1967; and US Department of Housing and Urban Development, 1968.

comprehend the location of one tract to another without the aid of supplementary information, typically in the form of maps. (Kao, 1963)

DBS accordingly formulated basic system objectives, of which the geographic aspects were:

(i) The geographical referencing of data observations within an ordered spatial universe or framework, the basic areal units of which would satisfy singly or in aggregation the area requirements of data users (subject to constraints of confidentiality and sampling and non-sampling errors associated with finely gauged areal units).

(ii) Ancillary data display capabilities; that is, graphic capabilities such as computer mapping that would derive from geographically referencing data by means of an automated spatial information system.

Several basic types of spatial framework were considered and, for urban areas, a street address conversion mode based on a 'nominal grid' and a 'geographic grid' was adopted for experimental development.(2)

1.2 The Purpose of the Study

The purpose of this study is to articulate the geographic basis of the DBS geocoding system for urban areas and to identify geographic considerations that ensue from system implementation. Of necessity the treatment is introductory.(3)

---

(2) Nominal grid denotes the street pattern and similar linear non-street features such as rivers, railways and boundaries. Geographic grid refers to a Cartesian grid having a known origin and scaled abscissa and ordinate axes that enable positions on the earth's surface to be reckoned in co-ordinate values.

(3) The study of the geographic basis of the DBS geocoding system for urban areas is largely an 'untilled field'. Though this paper seeks to reveal the basic elements in the concept that might be considered geographic in character, it is felt at the outset that more questions may be raised than will have been resolved satisfactorily.

Neither the procedures for implementing the system nor the benefits accruing through geographic referencing will be examined. Introductions to these topics can be found in: Canada. DBS, 1968; Fellegi and Weldon, 1967; and American Association of Geographers, 1964.

2. The Geographic Basis of the System

2.1 The Basic Concept

In order to furnish users with data for small areas specified by them on an 'ad hoc' basis, a spatial framework is required composed of primary areal units to which data can be referenced or coded (hence 'geocoding'), and from which, in aggregation, the query areas of data users can be constructed. Although the specific properties of the primary areal units will vary according to the spatial framework and system adopted, there are at least three significant characteristics that they must exhibit in a computerized system designed to retrieve data for user-specified areas; (i) they must be small enough to function as building-blocks; (ii) they must be identifiable by a code; and (iii) the location of each one must be unique and specific within an ordered spatial universe.

In urban areas DBS has approached the geographical referencing of data to primary areal units by means of a concept known generally as 'street address conversion'.(4) The logic behind street address conversion as a technique for geographical referencing runs as follows: The statistical population for which data are collected may be identified in urban areas by an address (specifically a civic, i.e. municipal, house number form of address, such as 1210 Carling Avenue, Ottawa, Ontario). Addresses are pre-grouped conveniently, in most cases,

(4) The street address conversion concept was introduced to the Dominion Bureau of Statistics by members of the Urban Data Center, University of Washington, Seattle. For the detailed development of the concept see Dial, 1964; Calkins, 1965; Crawford Jr., 1967.

into address ranges for block-faces(5) of city streets. A block-face is a meaningful areal unit to planners, administrators and researchers in general, and it is usually small enough that it could serve as the primary areal unit of a spatial framework. However, block-faces, while they may be a readily identifiable and discernible element of the street pattern, are not provided with a location-specific identification on the basis of their address ranges alone. The street name and address of a data observation do not describe the absolute location of the observation within an 'ordered' spatial framework, nor do they indicate relative location with respect to all other data observations.

In order for block-faces to function as primary areal units within an automated spatial information system they must each be provided with a location-specific identification. This latter step is achieved by relating the nominal grid to a geographic grid system (Vance, 1966; US Department of Housing and Urban Development, 1968). The Dominion Bureau of Statistics has adopted the 6° Universal Transverse Mercator Grid System for this purpose.

Block-faces may be represented as a point, the location of which in two-dimensional space is expressable as a set of co-ordinate values (x,y) within a geographic grid system. The co-ordinate identification of this point location is unique within the zone of the geographic grid system. The point that is chosen

(5) The term 'block-face' is used to designate one side of a street between neighbouring or consecutive intersections. The block-face constitutes the primary areal unit of the system's spatial framework, though on occasion it may be split either to respect the presence of a statistical area boundary or to retain some semblance of an optimum size. 'Address ranges' referred to in this study are the terminal civic house numbers for each block-face or split block-face. Their values are such that they enclose all individual civic house numbers in the primary areal unit.

is the mid-point of the long axis of the block-face set back from the street centre line a prescribed distance. This mid-point is termed the block-face 'centroid'.(6) The co-ordinate value for the centroid of the block-face can be used as a code that can be assigned to each data observation occurring on the block-face, thus identifying the data and simultaneously placing it in an ordered spatial framework.

A 'conversion' can be effected between the nominal grid elements (street name and address range by block-face) and the geographic grid element (the centroid of the block-face expressed in co-ordinates). Data bearing a street address identification can be tested by computer against a file of address ranges and corresponding centroid values. Once the appropriate address range has been found for the address in question, its centroid value can be substituted for the data's street address, and the data can be stored on the basis of this newly acquired identification. Geographical referencing or geocoding will have been accomplished.

2.2 Geographic Elements in the Concept

2.2.1 Spatial Framework The basic elements of the geographic framework of a street address conversion system are derived from the urban landscape and a geographic grid system. Only those features that would be vital to the development of a spatial framework for the referencing of data, and that would provide attendant plotting and graphic capabilities are abstracted from the urban landscape. It is with reference to these features that the term 'nominal grid'

(6) The 'centroid' is a point location situated at the mid-point of a block-face (or any other primary areal unit) and recessed a standard distance from the street centre line. Its co-ordinate values serve as both a code attributable to all data observations on the block-face, and as a unique location-specific identification for those data observations.

is used in this study. They may be classified as features of the street pattern and as non-street features, such as rivers, railways and area boundaries.

The spatial framework for the geographic referencing of data in urban areas is arrived at through the street pattern.(7) There are two basic elements: (i) portions of streets known as block-faces; and (ii) street address ranges for each of these (see footnote 5).

The nominal grid and the geographic grid are brought together by recording the mid-point or centroid of each block-face in terms of the co-ordinates of the geographic grid and relating all elements in a single address conversion file. For example:

Street name	Block-face address range	Co-ordinates of block-face mid-point (centroid)		
		(X)	(Y)	(Zone)
ADAM ST	1-19	481,209	4,896,212	12
	2-18	481,217	4,896,180	12

... etc.

The centroid code assigned to data observations by means of an address conversion file is the fundamental element of the automated system of geographic referencing of data for subsequent storage and retrieval.

(7) United States' plans along these lines have taken the form of Address Coding Guide (ACG) and Dual Independent Map Encoding (DIME) programmes for the 1970 Census.

2.2.2 <u>Plotting and Graphic Capabilities</u> The nominal grid lends itself to representation in a machine-processable form for both street features and non-street features and can, once associated with a geographic grid, be susceptible to automated map plotting and to computer mapping (Tobler, 1959).

Nominal grid features are represented in a machine-processable form by defining them in terms of the geographic grid. Each street feature and non-street feature, identified by its name, is coded as a string of points or 'nodes' that identify sequentially its terminals, intersections with other features and abrupt changes of direction. Each node, once identified by a geographic grid co-ordinate, has a location-specific identification that places the node uniquely within an ordered spatial universe and in relative position to all other points described therein. Automatic plotting of features is achieved by connecting node strings by straight lines on the basis of the node co-ordinate values.

The co-ordinate location of centroids is calculated using the geographic co-ordinate values of the nodes that define the block-faces; thus, both the establishment of the spatial framework (primary areal units, address ranges, centroids) and the definition of features are achieved by means of a mutual process.

The suitability of the geocoding system for general street mapping by computer at any specified scale has been demonstrated in programme development. The geocoded data base would also lend itself to line-printer and more sophisticated computer mapping routines.

3. Geographic Considerations Ensuing from Concept Implementation

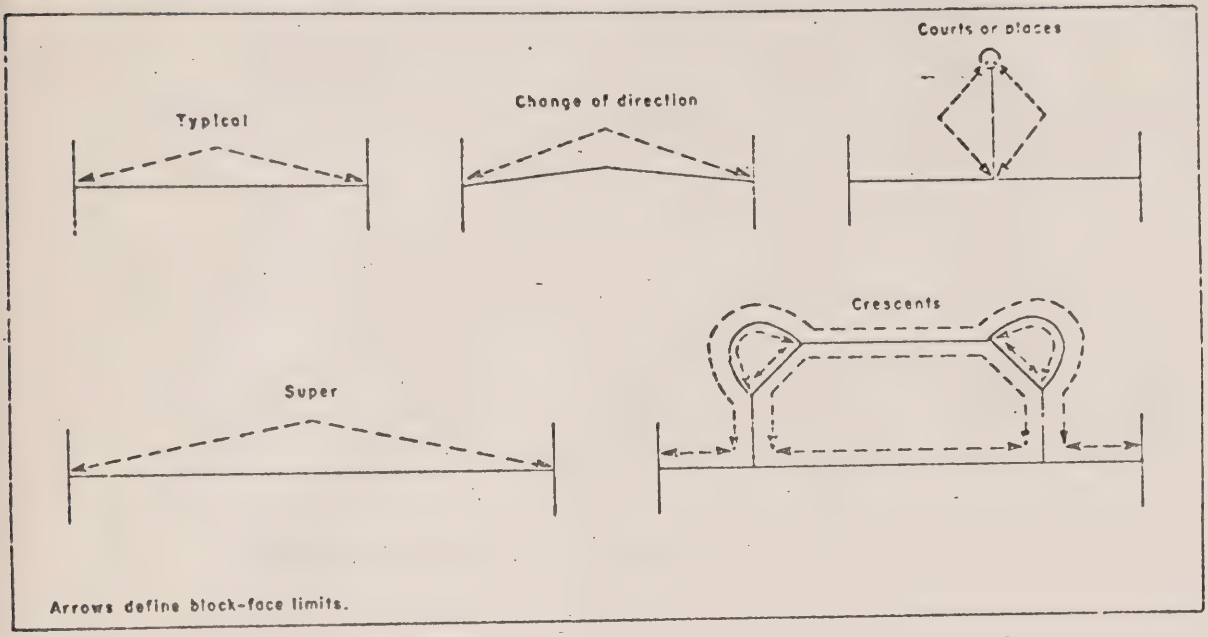
Each of the geographic elements in the street address conversion concept is examined below.

3.1 The Block-face

The block-face (see footnote 5) is one of two basic elements of the nominal grid; the other being the address range. The block-face was identified earlier in this paper as the primary areal unit of the spatial framework. Through this identification it is not difficult to understand how it has also been considered as the 'building-block' of the system, whereby through aggregation user-specified areas may be constructed. Conceptually, however, as will be seen below, the block-face is not so much the building-block of the system as is the 'address range'. Two basic functions of the block-face then become those of serving to contain physically the address range building-block, and to connote a convenient mental image or frame of reference with which the address range can be associated. Later in this paper the repercussions of choosing the block-face as the primary areal unit will be discussed in the context of data retrieval.

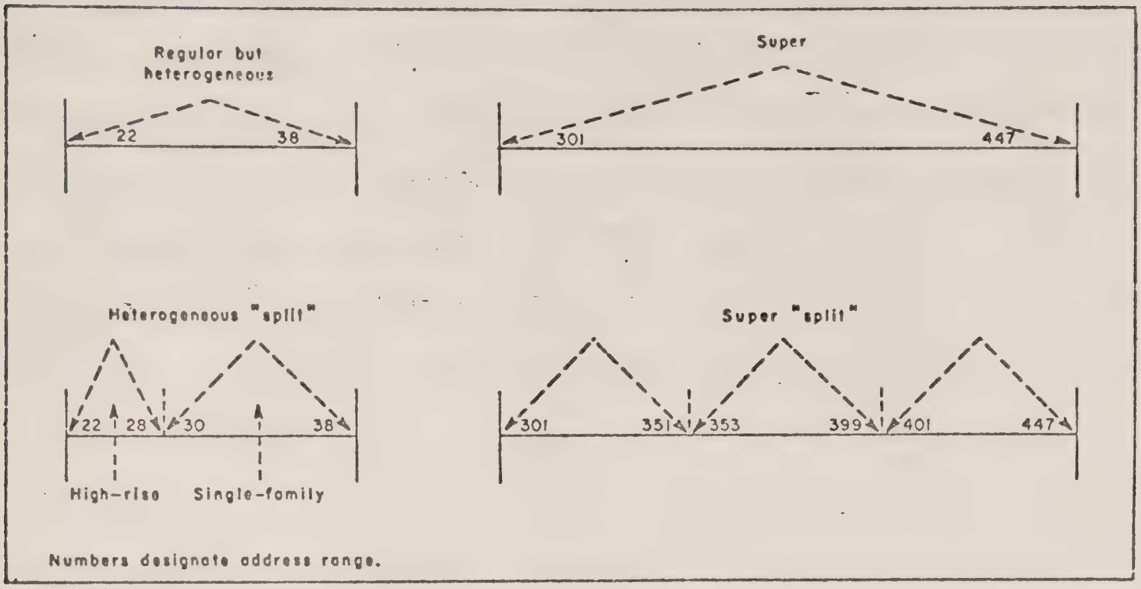
The block-face is perhaps most characterized by its lack of standardization. In this respect it is not at first glance a very satisfying prospect as a primary areal unit for a spatial framework. However, block-faces are generally small areal units that function adequately in conventional information systems and to which administrators, planners and presumably researchers at large can associate their information needs. In addition, the image of a block-face is readily projected, and actual block-faces can be delimited in the field. Figure 1 illustrates the variability in block-face dimensions and orientation.

FIGURE I. BLOCK-FACE DIMENSIONS AND ORIENTATION



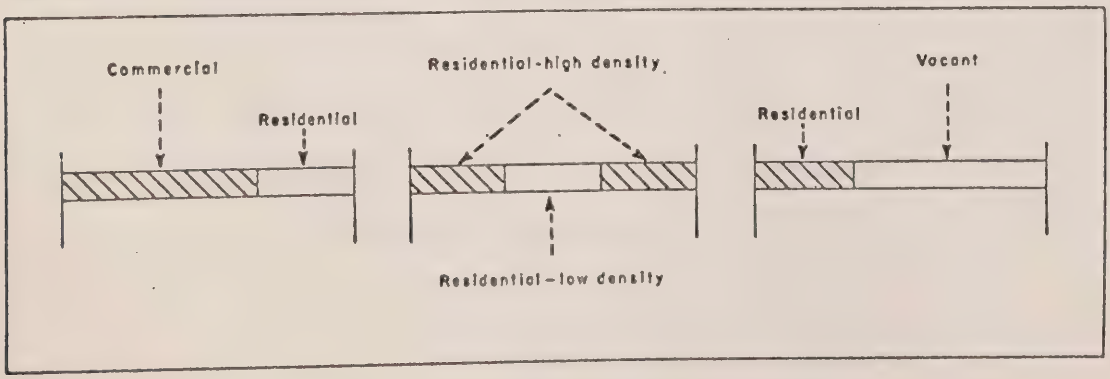
Conceptually, there is nothing sacrosanct about the block-face's integrity. A block-face may be "split" into two or more portions as long as the dividing line(s) is discernible on the nominal grid, and address ranges, which include the statistical population along each portion, may be allocated. The splitting of a block-face must be done before the centroid is calculated. A block-face may be segmented for several reasons, but generally to respect a statistical boundary or to divide a 'super block-face' into more typical-sized dimensions, or to separate a block-face of disparate characteristics into more homogeneous units. There exists, therefore, a measure of flexibility in primary areal unit creation (Fig. 2). The DBS system will split block-faces only in exceptional cases, notably of the super block-face variety, in an effort to preserve the spatial relationships of the nominal grid.

FIGURE 2. BLOCK-FACE SPLITS, INTO TWO OR MORE PRIMARY AREAL UNITS



Block-face variability is also present in terms of statistical population content. Though there may be a 'typical' block-face, statistically, many block-faces are heterogeneous in composition (Fig. 3). Nonetheless, it would appear that being small in size a block-face exhibits greater homogeneity within itself than it does with respect to a larger area such as a census tract of which it forms a part.

FIGURE 3. HETEROGENEITY OF BLOCK-FACE COMPOSITION



One concluding remark about the block-face should be made. The block-face is a significant entity within the street address conversion concept only in so far as its statistical population can be identified and expressed in terms of an address range. A block-face whose content is not expressed by an address range, notably vacant land or parks but possibly also dwellings, does not become a primary areal unit within the system. This situation can be rectified by establishing a pseudo-address range for the block-face and pseudo-addresses for the block-face components.

3.2 The Centroid

The centroid has been defined in footnote 6. Though the centroid is generally the mid-point of the block-face recessed a standard distance from the street centre line, it may also be the recessed mid-point of a segment of a split block-face. The centroid might better be referred to, conceptually at least, as the 'primary area data point' or perhaps 'coded data point'.

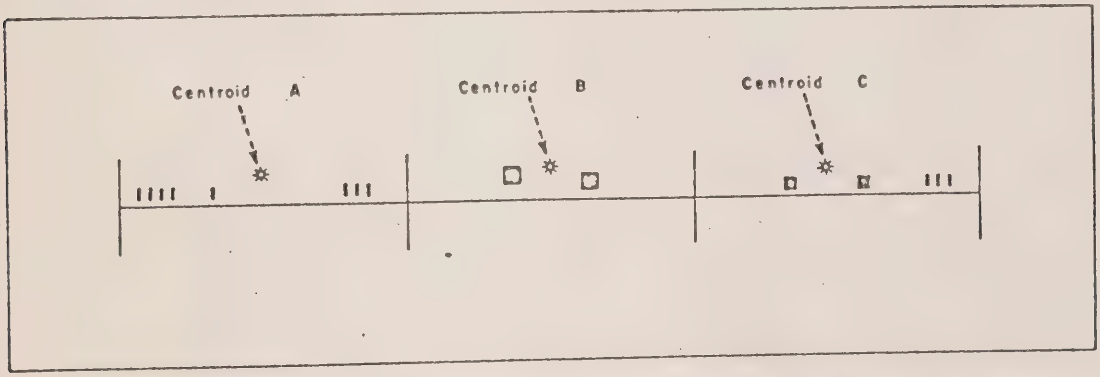
The nature of the centroid within the system is highly conditioned by the characteristics of the nominal grid, notably the street pattern and its block-face components. Some centroid characteristics are considered below.

The number of centroids is equal to the number of primary areal units (that is, block-faces and block-face segments) for which address ranges are present. The existing system, which calculates a centroid for a primary areal unit in association with an address range, could be modified to calculate a centroid for each and every block-face with or without addresses.

The centroid represents the final abstraction of the urban landscape into the context of the geographic grid, for purposes of geocoding. In effect "area" is mapped as a "point", leaving a vacuum surrounding discrete points in place of a continuous spatial surface or plane. Each centroid is a point, having no areal extent. The co-ordinate values of the centroid are simply part of the co-ordinate field of the geographic grid. Vance (1966, p. 31) points out that the accuracy of these co-ordinate values is a function of the map scale available. The centroid is linked to the primary areal unit (be it a block-face or a fraction of a block-face) by an address range expression of that unit.

Centroid data content, that is data attributable to the centroid for data storage, is variable, and for the same reasons mentioned in Section 3.1. As Figure 4 illustrates, the data content of centroids varies both in quantitative and qualitative terms.

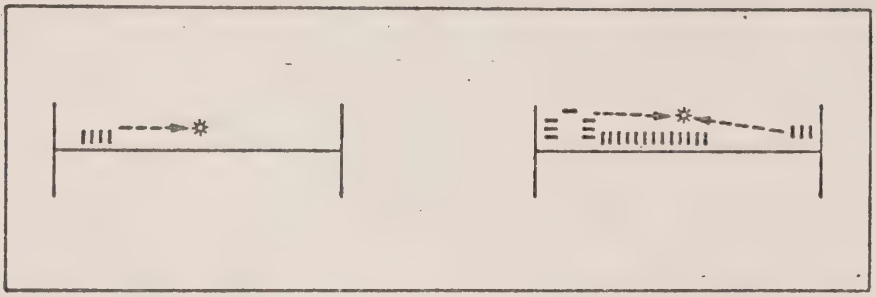
FIGURE 4. VARIATION IN CENTROID DATA CONTENT



In Figure 4 centroid 'A' is the code and spatial identification for 8 single-family dwellings; centroid 'B' groups two high-rise apartment buildings containing 200 dwellings; and lastly, centroid 'C' groups two small apartment buildings and three single-family houses for a total of 53 dwellings.

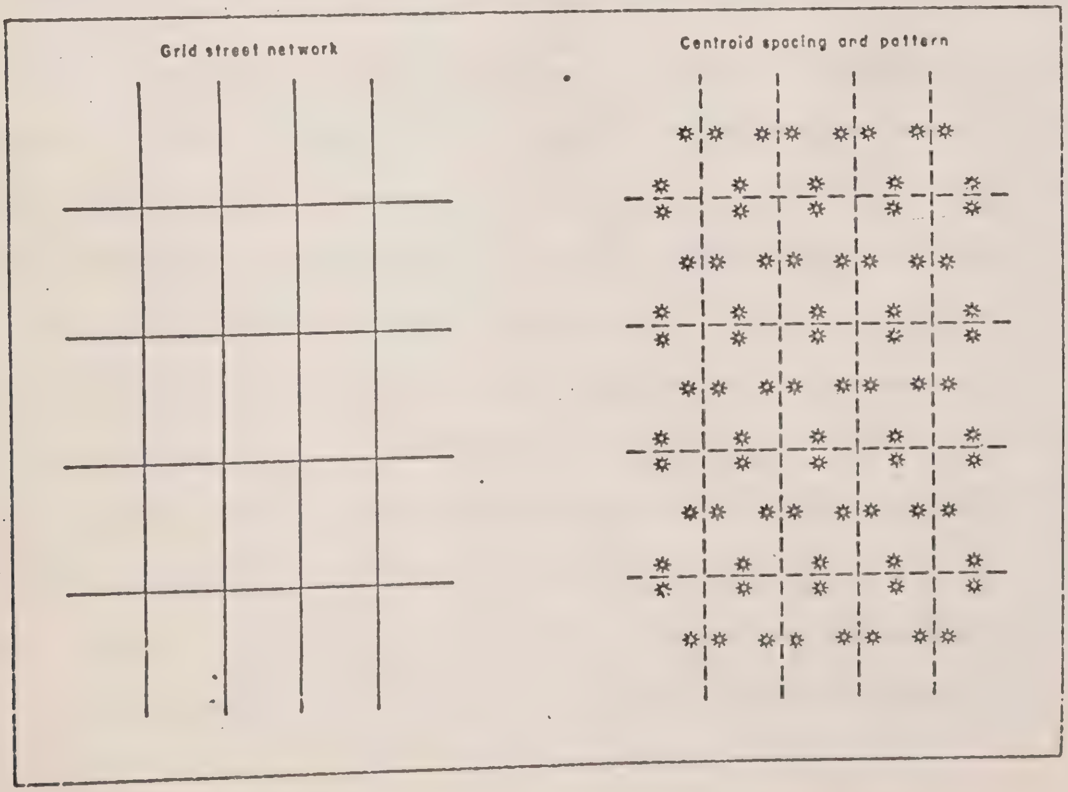
As mentioned above the centroid is the point representation of an area. The relocation of data from its absolute location to a representative point location follows logically. It is diagrammed below (Fig. 5) simply for emphasis, and because of its significance in a data users' context.

FIGURE 5. DATA RELOCATION TO THE CENTROID



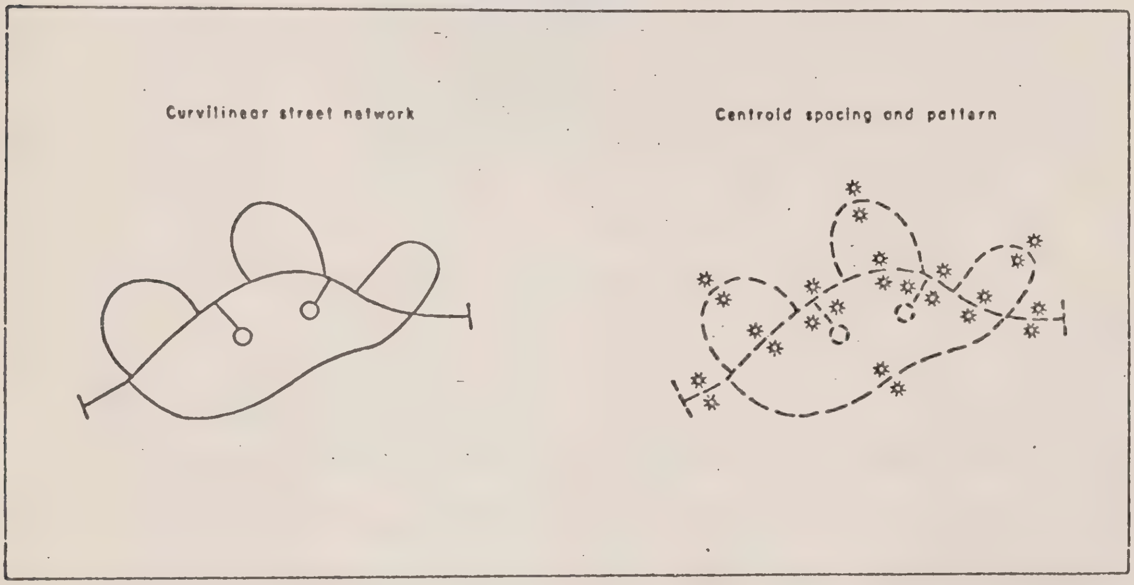
The spacing of centroids reflects naturally the geometry of the nominal grid, specifically the street pattern. A regular grid street pattern (assuming address ranges throughout) results in a regular, orderly centroid distribution (see Fig. 6).

FIGURE 6. CENTROID SPACING AND PATTERN IN A REGULAR GRID STREET NETWORK



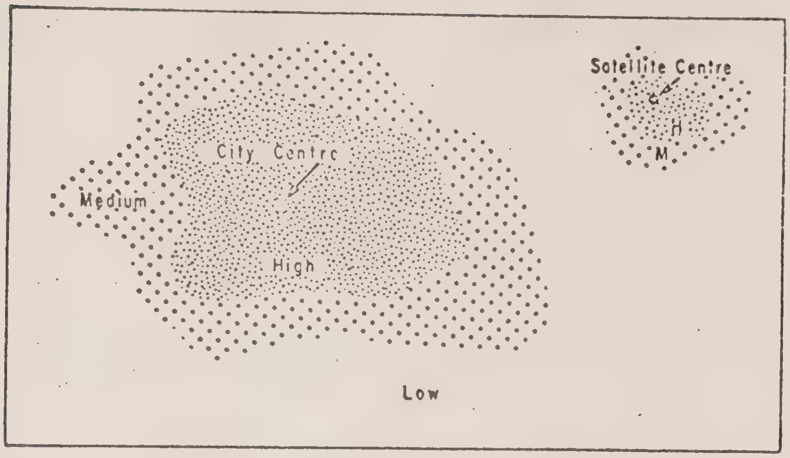
In less regular street pattern areas the resultant spacing of centroids is less orderly. An example of the distribution of centroids in a curvilinear street area (a type found increasingly in suburban areas and new towns) is shown in Figure 7.

FIGURE 7. CENTROID SPACING AND PATTERN IN A CURVILINEAR STREET NETWORK



The density of centroids, that is, the number of centroids per specified areal unit, is not constant, but does appear to exhibit certain tendencies of a locational nature. A cursory examination of centroids in one city alone, London, Ontario, revealed centroid density to be higher in the older central part of the city, where centroids numbered some 100-120 per square kilometre, than in the surrounding built-up residential areas, where some 60-80 centroids per square kilometre were calculated. Conceptually, therefore, centroid density diminishes with distance from the centre of a city and with the progression from an ordered grid street network to an increasingly curvilinear and random street pattern. This hypothesis, however, warrants further research in view of its implications to data users. An attempt is made in Figure 8 to illustrate the hypothesis.

FIGURE 8. A HYPOTHETICAL CENTROID DENSITY IN URBAN AREAS

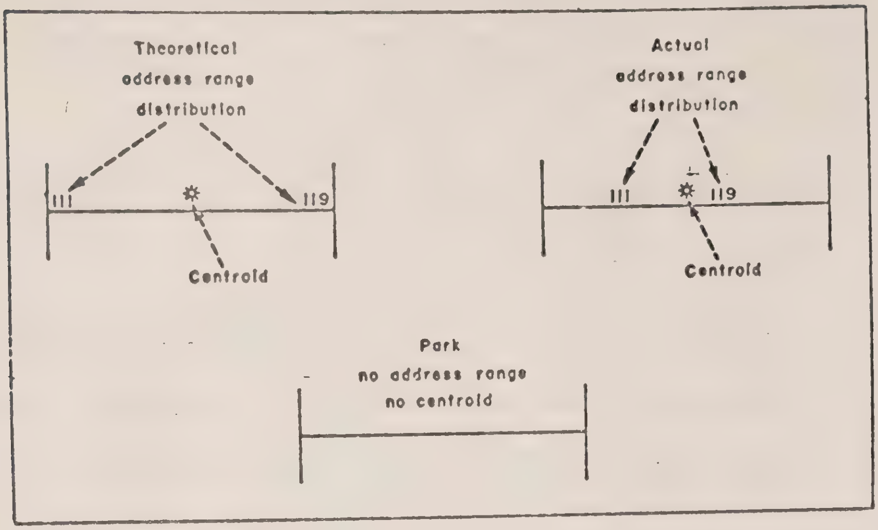


3.3 The Address Range

The address range is defined in footnote 5. In the context of a street address conversion system address ranges form the fundamental link between the nominal grid elements — the primary areal units (block-faces and fractions of block-faces) and the geographic grid elements — the centroids.

The address range effectively defines the statistical population that will be ascribed to the centroid. It exercises in this capacity a discriminating function, permitting only address-bearing data to enter the system. It is considered to enclose within its values all of the address numbers of a primary areal unit, but nothing is said about the location or distribution of individual addresses within that areal unit. For this reason and the fact that many block-faces do not contain address ranges, it is more accurate to consider the address range (as opposed to the block-face) as the true 'building-block' of the system (see Fig. 9).

FIGURE 9. THE ADDRESS RANGE IN TERMS OF THE BLOCK-FACE

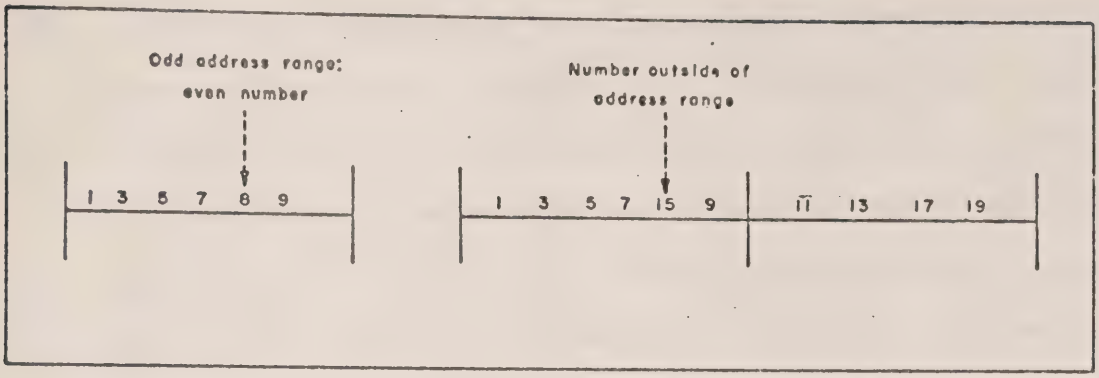


Although forming a viable link between the primary areal units and the geographic grid, the use of the address range is not without certain disadvantages, most of which can be overcome.

In the first place it is necessary to define an address range for each primary areal unit. Once obtained such address ranges must be maintained up-to-date, reflecting changes in the street and addressing pattern as the city develops.

A second difficulty arises from anomalies in addressing systems. Though most civic house numbers increase sequentially along a block-face, the system must contend with even numbers occurring within odd address ranges (and vice versa), and occasionally with civic house numbers that cannot be included within a given address range (see Fig. 10).

FIGURE 10. ANOMALIES IN ADDRESS SYSTEMS

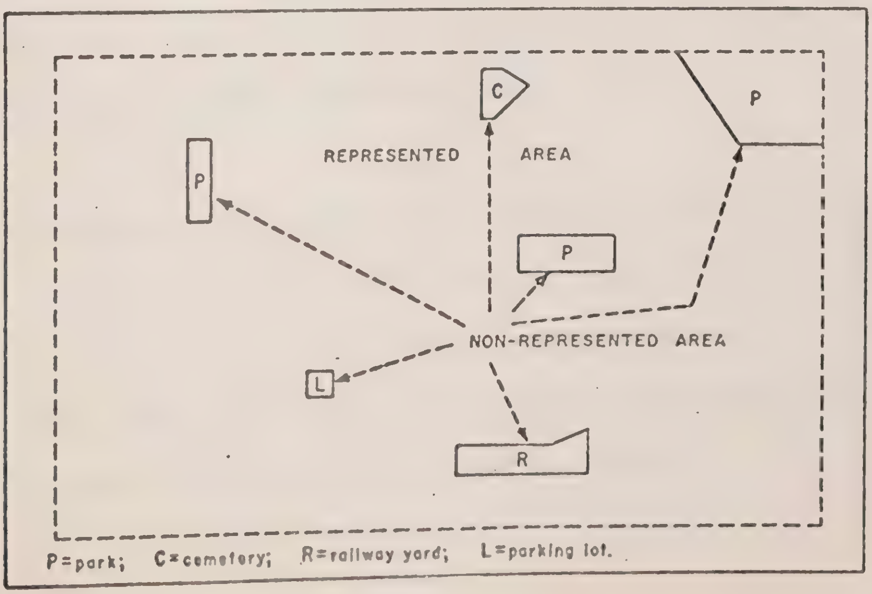


No difficulty occurs if the addresses are out of sequence, as long as they are either all odd or all even and the terminal addresses contain the others.

A third difficulty presents itself in the case of dwellings (or any other statistical population) that do not bear an address. Such a situation, it is felt, is infrequent in large urban areas, and can be resolved by providing a 'pseudo' address for each occurrence.

The fact that centroids are only calculated for primary areal units with which address ranges are associated categorically means that much of an urban area is not represented within the system (except for plotting purposes). These non-represented areas, such as parks, railway yards, cemeteries, parking lots, may form a significant part of the actual urban 'area' (see Fig. 11).

FIGURE 11. REPRESENTED VERSUS ACTUAL URBAN AREA



The disadvantages, mostly operational in nature, of the address range are minor in comparison with the major advantage of its widespread use. More specifically, individual addresses are used extensively as identification information for statistical populations by many disparate governmental and private agencies and firms. Since any data identified by an address can be accorded a centroid value, the data storage design can be 'open-ended' to include additional survey data.

4. Geographic Aspects of Data Storage and Retrieval

A geocoded census data base may be viewed conceptually as one large storage file, though this file may be split into sub-files.

The data observations derived from the census questionnaire will be kept separate for each member of the statistical population. The members of the statistical population will be arranged within the geocoded census data file in terms of their position within their primary areal unit. Though kept discrete for purposes of cross-tabulation of characteristics, all members of the statistical population (and their associated characteristics) are 'grouped' under the centroid of the primary areal units in which they are situated. The orderly arrangement of the primary areal units themselves within the census data file will be according to the co-ordinate values of their respective centroids.

As noted above, the geocoded census data file can be viewed conceptually as one file. Since this file is to be structured geographically on the basis of the co-ordinate values of centroids, it will merge the primary areal units determined under the urban system of geocoding with those primary areal units (of a different nature) designated within the rural geocoding system.

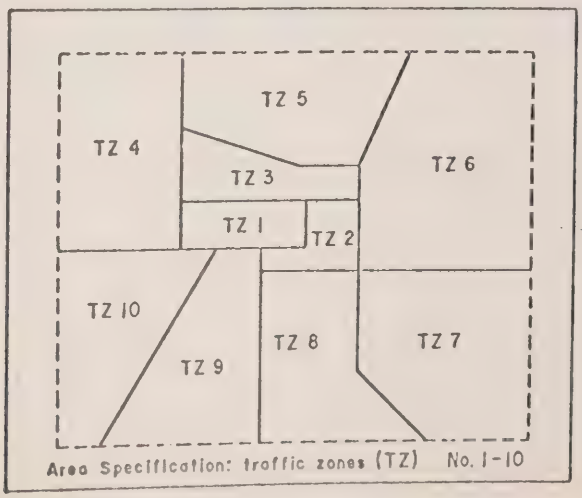
The ordering of primary areal units within the geocoded census data file on the basis of their 'centroids' co-ordinate values has advantages at the data retrieval stage. Some geographic considerations concerning retrieval from a geocoded data base are identified below.

Any request for data from users must incorporate four elements: statistical population, variable(s), time and space. The geocoding system has been designed primarily to contend with the spatial dimension discernible in any user request. In a narrower sense it has been designed to allow for data retrieval by user-specified areas of sub-municipality size. The system seeks to provide tabulations for 'ad hoc' areas specified by users by manipulating data for finely-gauged primary areal units that will aggregate as closely as possible to form the requested areas.

User-specified areas are expected to fall into four main categories: (i) bounded areas or 'polygons'; (ii) concentric zones or distance bands; (iii) street-oriented areas and, (iv) uniform data regions.

Figure 12 gives an example of a polygon request in which the 'area specification' called for data tabulation by traffic zones No. 1-10 for a certain area.

FIGURE 12. USER-SPECIFIED AREAS: POLYGONAL



Other examples of bounded polygon sets might include: school districts, planning neighbourhoods, wards, police districts.

Figure 13 illustrates a request for data tabulations by concentric zones, as defined by radii from a central point.

FIGURE 13. USER-SPECIFIED AREAS: CONCENTRIC

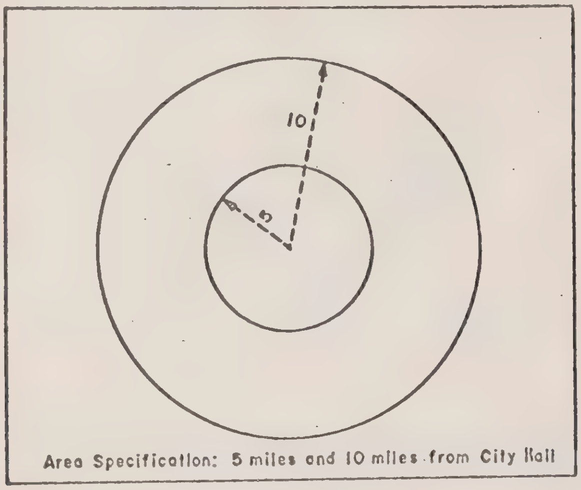
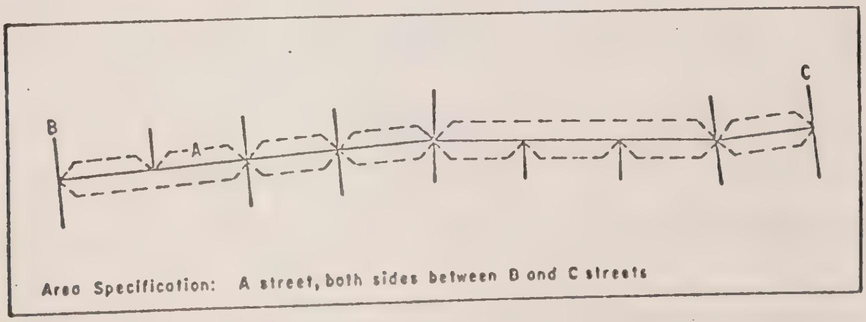


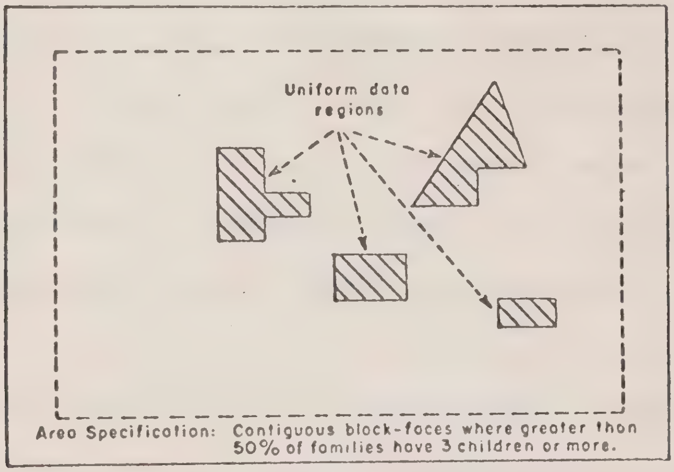
Figure 14 illustrates the specifications of a request for data for both sides of a street between specified intersections.

FIGURE 14. USER-SPECIFIED AREAS: STREET-ORIENTED



(iv) Figure 15 illustrates a form of areal retrieval that may ultimately be practicable; i.e. retrieval of tabulated data and regional delimitation based on homogeneous characteristics. This form of retrieval, the opposite of predefined areas, is discussed briefly by Fellegi and Weldon 1967, p. 55.

FIGURE 15. USER-SPECIFIED AREAS: UNIFORM DATA



The suitability of the block-face and split block-face as primary areal units to construct user-specified areas will vary according to the type of area specification submitted by the user.

The block-face will be a suitable primary areal unit in the case of bounded user request areas, type (i) above, as long as the boundaries of the request areas do not dissect block-faces. If these boundaries cut block-faces in places, then the tabulation can only be considered approximate.

In type (ii) above, concentric zones, the boundary of the user request area(s) is a circle, and hence tabulations from block-face primary areal units will of necessity be approximate. Requests of this nature generally seek only an approximate answer.

When users request data for street-oriented areas, type (iii) above, block-face aggregation will provide an exact answer.

Similarly, the request for data by uniform data regions (inverse retrieval) acknowledges from the outset the suitability of the block-face primary areal unit — since no finer area unit exists to function in that capacity.

The general approach to data retrieval defines the boundaries of user areas in terms of the geographic grid, and subsequently searches the file of centroids and, using a point-in-polygon geometric algorithm, determines whether these centroids according to their co-ordinate values are 'inside' or 'outside' the defined areas. The data coded with the centroids found to lie inside the defined areas are tabulated in accordance with the user's specifications.

One final note on data retrieval should be made. Primary areal units, being small and containing a variable statistical population content may present problems, even in aggregation, in terms of maintaining confidentiality and/or of exhibiting a high sampling and non-sampling error. In the former case, suppression of data would be required in accordance with the regulations of the Statistics Act and, in the latter case, an estimate of such errors might have to be provided with the tabulations.

5. Summary and Conclusion

Those elements of the street address conversion mode of geocoding for urban areas that might be considered as geographic in nature have been identified as the block-face, the centroid and the address range. The block-face and the address range have been construed as elements of a nominal grid that functions as a spatial framework for referencing data. However, as the nominal grid is an imperfect one with which to determine either relative or absolute location, it is merged with a geographic grid system by providing each primary areal unit of the nominal grid with a centroid, or co-ordinate location within the geographic grid.

The block-face, block-face centroid and block-face address range elements have each been examined in detail in order to describe the role that they play in a street address conversion system and to determine the manner in which their characteristics effect the attainment of geocoding system objectives. The study reveals that the address range may be considered the real building-block of the spatial framework; though the block-face serves to physically contain this range and to convey a convenient mental image of it. The centroid, however, is the element most fundamental to system operation. Expressed in geographic co-ordinates, it serves as a code attributable to data observations for the block-face in question and as a unique location-specific identifier for those observations. The Cartesian relationship of centroid co-ordinates permits data for all block-faces to be stored in a single randomly accessible data base, for subsequent aggregation by block-faces to satisfy user requests for data by non-standard areas.

The limitations of each element have also been pointed out. None of these, however, would seriously curtail system objectives.

The following point can be made as a result of this study. The geographic elements of the system have been shown to underpin the entire system concept. It seems reasonable to conclude that the development of any spatial information system should include fundamental research into the nature of the spatial framework that ultimately will affect the utility of the system.

References

American Association of Geographers, Census Advisory Committee. "Geographic Ordering of Information: New Opportunities". The Professional Geographer, 16 (4, 1964), 39-44.

Calkins, H.W. "Operations Manual for Street Address Conversion System". Research Report No. 2. Seattle: Urban Data Center, University of Washington, 1965.

Canada. Dominion Bureau of Statistics. Geographically Referenced Data Storage and Retrieval System. Sampling and Survey Research Staff. Ottawa, 1968 (Processed).

Crawford Jr., R.J. "Utility of an Automated Geocoding System for Urban Land Use Analysis". Research Report No. 3. Seattle: Urban Data Center, University of Washington, 1967.

Dial, R.B. "Street Address Conversion System". Research Report No. 1. Seattle: Urban Data Center, University of Washington, 1964.

Fellegi, I.P., and Weldon, J.I. "Computer Methods of Geographical Coding and Retrieval of Data in the Dominion Bureau of Statistics, Canada". Proceedings of the Social Statistics Section, American Statistical Association, Washington (1967), 53-57.

Hagerstrand, T. "Statistika primaruppgifter, flygkartering och "Data Processing" — maskiner: Ett kombinerings-projekt". Svensk Geogr. Arsbok, 31 (1955), 254.

Haggett, Peter. Locational Analysis in Human Geography. London: Edward Arnold (Publishers) Ltd., 1965.

Kao, Richard C. "The Use of Computers in the Processing and Analysis of Geographic Information". The Geographical Review, 53 (1963), 530-547.

Simmons, James W. "Urban Geography in Canada". The Canadian Geographer, II (Winter, 1967), 341-356.

Tobler, Waldo R. "Automation and Cartography". The Geographical Review, 49 (October 1959), 526-534.

US Department of Housing and Urban Development. Urban and Regional Information Systems: Support for Planning in Metropolitan Areas. Washington, D.C., US Government Printing Office, 1968.

Vance, J.A. "Geographical Data Coding Grid". Paper presented at the 47th Annual Convention of the Canadian Good Roads Association. Halifax, N.S., 1966 (Processed).